(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 373 203 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication of patent specification: **31.08.94** (51) Int. Cl.5: **C12Q 1/68**

(21) Application number: **89905449.8**

(22) Date of filing: **02.05.89**

(86) International application number:
PCT/GB89/00460

(87) International publication number:
WO 89/10977 (16.11.89 89/27)

(54) Method and apparatus for analysing polynucleotide sequences.

(30) Priority: **03.05.88 GB 8810400**

(43) Date of publication of application:
20.06.90 Bulletin 90/25

(45) Publication of the grant of the patent:
31.08.94 Bulletin 94/35

(84) Designated Contracting States:
AT BE CH DE FR GB IT LI LU NL SE

(56) References cited:
EP-A- 228 075
WO-A-85/01051
GB-A- 2 156 074

(73) Proprietor: **ISIS INNOVATION LIMITED**
2 South Parks Road
Oxford OX1 3UB (GB)

(72) Inventor: **SOUTHERN, Edwin**
30 Staverton Road
Oxford OX2 6XJ (GB)

(74) Representative: **Pennant, Pyers et al**
Stevens, Hewlett & Perkins
1 Serjeants' Inn
Fleet Street
London EC4Y 1LL (GB)

EP 0 373 203 B1

Description

1. INTRODUCTION

This invention provides a method and apparatus for analysing a polynucleotide sequence, either an unknown sequence or a known sequence. A support, e.g. a glass plate, carries an array of the whole or a chosen part of a complete set of oligonucleotides, the different oligonucleotides occupying separate cells of the array and being capable of taking part in hybridisation reactions. The array may comprise one or more pairs of oligonucleotides. The polynucleotide sequence, or fragments thereof, are labelled and applied to the array under hybridising conditions. Applications include analysis of known point mutations, genomic fingerprinting, linkage analysis, characterisation of mRNAs, mRNA populations, and sequence determination.

Three methods dominate molecular analysis of nucleic acid sequences: gel electrophoresis of restriction fragments, molecular hybridisation, and the rapid DNA sequencing methods. These three methods have a very wide range of applications in biology, both in basic studies, and in the applied areas of the subject such as medicine and agriculture. Some idea of the scale on which the methods are now used is given by the rate of accumulation of DNA sequences, which is now well over one million base pairs a year. However, powerful as they are, they have their limitations. The restriction fragment and hybridisation methods give a coarse analysis of an extensive region, but are rapid; sequence analysis gives the ultimate resolution, but it is slow, analysing only a short stretch at a time. There is a need for methods which are faster than the present methods, and in particular for methods which cover a large amount of sequence in each analysis.

This invention provides a new approach which produces both a fingerprint and a partial or complete sequence in a single analysis, and may be used directly with complex DNAs and populations of RNA without the need for cloning.

In one aspect the invention provides a method of analysing a polynucleotide sequence, by the use of a support to the surface of which is attached an array of the whole or a chosen part of a complete set of oligonucleotides of chosen lengths, the different oligonucleotides occupying separate cells of the array, which method comprises labelling the polynucleotide sequence or fragments thereof, applying the polynucleotide sequence or fragments thereof under hybridisation conditions to the array, and observing the location of the label on the surface associated with particular members of the set of oligonucleotides.

In another aspect, the invention provides apparatus suitable for analysing a polynucleotide sequence by the above method, comprising a support and attached to a surface thereof, an array of the whole or a chosen part of a complete set of oligonucleotides of chosen lengths the different oligonucleotides occupying separate cells of the array and being capable of taking part in hybridisation reactions.

In another aspect, the invention provides a method for generating an array of oligonucleotides within discrete cells of a support material comprising the steps of

a) segregating a support material into discrete cell locations;
b) coupling a nucleotide to a first set of cell locations;
c) coupling a nucleotide to a second set of cell locations;
d) coupling a nucleotide to a third set of cell locations;
e) and continuing the sequence of coupling steps until the desired array has been generated,

the coupling being effected at each location either to the surface of the support or to a nucleotide coupled in a previous step at that location.

The oligonucleotides of the array are preferably of defined sequence, and preferably bound to the support by a covalent link through a terminal nucleotide.

The idea of the invention is thus to provide a structured array of the whole or a chosen part of a complete set of oligonucleotides of one or several chosen lengths. The array, which may be laid out on a supporting film or glass plate, forms the target for a hybridisation reaction. The chosen conditions of hybridisation and the length of the oligonucleotides must at all events be sufficient for the available equipment to be able to discriminate between exactly matched and mismatched oligonucleotides. In the hybridisation reaction, the array is explored by a labelled probe, which may comprise oligomers of the chosen length or longer polynucleotide sequences or fragments, and whose nature depends on the particular application. For example, the probe may comprise labelled sequences amplified from genomic DNA by the polymerase chain reaction, or a mRNA population, or a complete set of oligonucleotides from a complex sequence such as an entire genome. The end result is a set of filled cells corresponding to the oligonucleotides present in the analysed sequence, and a set of "empty" sites corresponding to the sequences which are absent in the analysed sequence. The pattern produces a fingerprint representing all of the sequence analysed. In addition, it is possible to assemble most or all of the sequence analysed if an

oligonucleotide length is chosen such that most or all oligonucleotide sequences occur only once.

The number, the length and the sequences of the oligonucleotides present in the array "lookup table" also depend on the application. The array may include all possible oligonucleotides of the chosen length, as would be required if there was no sequence information on the sequence to be analysed. In this case, the preferred length of oligonucleotide used depends on the length of the sequence to be analysed, and is such that there is likely to be only one copy of any particular oligomer in the sequence to be analysed. Such arrays are large. If there is any information available on the sequence to be analysed, the array may be a selected subset. For the analysis of a sequence which is known, the size of the array is of the same order as length of the sequence, and for many applications, such as the analysis of a gene for mutations, it can be quite small. These factors are discussed in detail in what follows.

2. OLIGONUCLEOTIDES AS SEQUENCE PROBES

Oligonucleotides form base paired duplexes with oligonucleotides which have the complementary base sequence. The stability of the duplex is dependent on the length of the oligonucleotides and on base composition. Effects of base composition on duplex stability can be greatly reduced by the presence of high concentrations of quaternary or tertiary amines. However, there is a strong effect of mismatches in the oligonucleotides duplex on the thermal stability of the hybrid, and it is this which makes the technique of hybridisation with oligonucleotides such a powerful method for the analysis of mutations, and for the selection of specific sequences for amplification by DNA polymerase chain reaction. The position of the mismatch affects the degree of destabilisation. Mismatches in the centre of the duplex may cause a lowering of the Tm by 10°C compared with 1°C for a terminal mismatch. There is then a range of discriminating power depending on the position of mismatch, which has implications for the method described here. There are ways of improving the discriminating power, for example by carrying out hybridisation close to the Tm of the duplex to reduce the rate of formation of mismatched duplexes, and by increasing the length of oligonucleotide beyond what is required for unique representation. A way of doing this systematically is discussed.

3. ANALYSIS OF A PREDETERMINED SEQUENCE

One of the most powerful uses of oligonuoleotide probes has been in the detection of single base changes in human genes. The first example was the detection of the single base change in the betaglobin gene which leads to sickle cell disease. There is a need to extend this approach to genes in which there may be a number of different mutations leading to the same phenotype, for example the DMD gene and the HPRT gene, and to find an efficient way of scanning the human genome for mutations in regions which have been shown by linkage analysis to contain a disease locus for example Huntington's disease and Cystic Fibrosis. Any known sequence can be represented completely as a set of overlapping oligonucleotides. The size of the set is s + 1 = N, where N is the length of the sequence and s is the length of an oligomer. A gene of 1 kb for example, may be divided into an overlapping set of around one thousand oligonucleotides of any chosen length. An array constructed with each of these oligonucleotides in a separate cell can be used as a multiple hybridisation probe to examine the homologous sequence in any context, a single-copy gene in the human genome or a messenger RNA among a mixed RNA population, for example. The length s may be chosen such that there is only a small probability that any oligomer in the sequence is represented elsewhere in the sequence to be analysed. This can be estimated from the expression given in the section discussing statistics below. For a less complete analysis it would be possible to reduce the size of the array e.g. by a factor of up to 5 by representing the sequence in a partly or non-overlapping set. The advantage of using a completely overlapping set is that it provides a more precise location of any sequence difference, as the mismatch will scan in s consecutive oligonucleotides.

4. ANALYSIS OF AN UNDETERMINED SEQUENCE

The genomes of all free living organisms are larger than a million base pairs and none has yet been sequenced completely. Restriction site mapping reveals only a small part of the sequence, and can detect only a small portion of mutations when used to compare two genomes. More efficient methods for analysing complex sequences are needed to bring the full power of molecular genetics to bear on the many biological problems for which there is no direct access to the gene or genes involved. In many cases, the full sequence of the nucleic acids need not be determined; the important sequences are those which differ between two nucleic acids. To give three examples: the DNA sequences which are different between a wild

type organism and one which carries a mutant can lead the way to isolation of the relevant gene; similarly, the sequence differences between a cancer cell and its normal counterpart can reveal the cause of transformation; and the RNA sequences which differ between two cell types point to the functions which distinguish them. These problems can be opened to molecular analysis by a method which identifies sequence differences. Using the approach outlined here, such differences can be revealed by hybridising the two nucleic acids, for example the genomic DNA of the two genotypes, or the mRNA populations of two cell types to an array of oligonucleotides which represent all possible sequences. Positions in the array which are occupied by one sequence but not by the other show differences in two sequences. This gives the sequence information needed to synthesise probes which can then be used to isolate clones of the sequence involved.

4.1 ASSEMBLING THE SEQUENCE INFORMATION

Sequences can be reconstructed by examining the result of hybridisation to an array. Any oligonucleotide of length s from within a long sequence, overlaps with two others over a length s-1. Starting from each positive oligonucleotide, the array may be examined for the four oligonucleotides to the left and the four to the right that can overlap with a one base displacement. If only one of these four oligonucleotides is found to be positive to the right, then the overlap and the additional base to the right determine s bases in the unknown sequence. The process is repeated in both directions, seeking unique matches with other positive oligonucleotides in the array. Each unique match adds a base to the reconstructed sequence.

4.2 SOME STATISTICS

Any sequence of length N can be broken down to a set of ~ N overlapping sequences s base pairs in length. (For double stranded nucleic acids, the sequence complexity of a sequence of N base pairs is 2N, because the two strands have different sequences, but for the present purpose, this factor of two is not significant). For oligonucleotides of length s, there are 4^s different sequence combinations. How big should s be to ensure that most oligonucleotides will be represented only once in the sequence to be analysed, of complexity N? For a random sequence the expected number of s-mers which will be present in more than one copy is

$$\mu_{>1} \approx 4^s(1-e^{-\lambda}(1+\lambda))$$

where

$$\lambda = (N - s + 1)/4^s$$

For practical reasons it is also useful to know how many sequences are related to any given s-mer by a single base change. Each position can be substituted by one of three bases, there are therefore 3s sequences related to an individual s-mer by a single base change, and the probability that any s-mer in a sequence of N bases is related to any other s-mer in that sequence allowing one substitution is $3s \times N/4^s$. The relative signals of matched and mismatched sequences will then depend on how good the hybridisation conditions are in distinguishing a perfect match from one which differ by a single base. (If 4^s is an order of magnitude greater than N, there should only be a few, 3s/10, related to any oligonucleotide by one base change.) The indications are that the yield of hybrid from the mismatched sequence is a fraction of that formed by the perfect duplex.

For what follows, it is assumed that conditions can be found which allow oligonucleotides which have complements in the probe to be distinguished from those which do not.

4.3 ARRAY FORMAT, CONSTRUCTION AND SIZE

To form an idea of the scale of the arrays needed to analyse sequences of different complexity it is convenient to think of the array as a square matrix. All sequences of a given length can be represented just once in a matrix constructed by drawing four rows representing the four bases, followed by four similar columns. This produces a 4 x 4 matrix in which each of the 16 squares represents one of the 16 doublets. Four similar matrices, but one quarter the size, are then drawn within each of the original squares. This produces a 16 x 16 matrix containing all 256 tetranucleotide sequences. Repeating this process produces a

matrix of any chosen depth, s, with a number of cells equal to 4^s. As discussed above, the choice of s is of great importance, as it determines the complexity of the sequence representation. As discussed below, s also determines the size of the matrix constructed, which must be very big for complex genomes. Finally, the length of the oligonucleotides determines the hybridisation conditions and their discriminating power as hybridisation probes.

s	4^s	Genomes	Side of Matrix (pixel = 100 μm)	Number of Sheets of film
8	65536	$4^s x^{10}$		
9	262144			
10	1.0×10^6	cosmid	100 mm	1
11	4.2×10^6			
12	1.7×10^7			
13	6.7×10^7	E.coli		
14	2.6×10^8	yeast	1.6 m	9
15	1.1×10^9			
16	4.2×10^9			
17	1.7×10^{10}			
18	6.7×10^{10}	human	25 m	2,500
19	2.7×10^{11}			
20	1.1×10^{12}		100 m	

The table shows the expected scale of the arrays needed to perform the first analysis of a few genomes. The examples were chosen because they are genomes which have either been sequenced by conventional procedures - the cosmid scale -, are in the process of being sequenced - the E. coli scale -, or for which there has been considerable discussion of the magnitude of the problem - the human scale. The table shows that the expected scale of the matrix approach is only a small fraction of the conventional approach. This is readily seen in the area of X-ray film that would be consumed. It is also evident that the time taken for the analysis would be only a small fraction of that needed for gel methods. The "Genomes" column shows the length of random sequence which would fill about 5% of cells in the matrix. This has been determined to be the optimum condition for the first step in the sequencing strategy discussed below. At this size, a high proportion of the positive signals would represent single occurrences of each oligomer, the conditions needed to compare two genomes for sequence differences.

5. REFINEMENT OF AN INCOMPLETE SEQUENCE

Reconstruction of a complex sequence produces a result in which the reconstructed sequence is interrupted at any point where an oligomer that is repeated in the sequence occurs. Some repeats are present as components of long repeating structures which form part of the structural organisation of the DNA, dispersed and tandem repeats in human DNA for example. But when the length of oligonucleotide used in the matrix is smaller than that needed to give totally unique sequence representation, repeats occur by chance. Such repeats are likely to be isolated. That is, the sequences surrounding the repeated oligomers are unrelated to each other. The gaps caused by these repeats can be removed by extending the sequence to longer oligomers. In principle, those sequences shown to be repeated by the first analysis; using an array representation of all possible oligomers, could be resynthesised with an extension at each end. For each repeated oligomer, there would be 4 x 4 = 16 oligomers in the new matrix. The hybridisation analysis would now be repeated until the sequence was complete. In practice, because the results of a positive signal in the hybridisation may be ambiguous, it may be better to adopt a refinement of the first result by extending all sequences which did not give a clear negative result in the first analysis. An advantage of this approach is that extending the sequence brings mismatches which are close to the ends in the shorter oligomer, closer to the centre in the extended oligomer, increasing the discriminatory power of duplex formation.

5.1 A HYPOTHETICAL ANALYSIS OF THE SEQUENCE OF BACTERIOPHAGE λ DNA

Lambda phage DNA is 48,502 base pairs long. Its sequence has been completely determined, we have treated one strand of this as a test case in a computer simulation of the analysis. The table shows that the appropriate size of oligomer to use for a sequence of this complexity is the 10-mer. With a matrix of 10-

mers, the size was 1024 lines square. After "hybridisation" of the lambda 10-mers in the computer, 46,377 cells were positive, 1957 had double occurrences, 75 triple occurrences, and three quadruple occurrences. These 46,377 positive cells represented known sequences, determined from their position in the matrix. Each was extended by four x one base at the 3' end and four x one base at the 5', end to give 16 x 46,377 = 742,032 cells. This extended set reduced the number of double occurrences to 161, a further 16-fold extension brought the number down to 10, and one more provided a completely overlapped result. Of course, the same end result of a fully overlapped sequence could be achieved starting with a 4^{16} matrix, but the matrix would be 4000 times bigger than the matrix needed to represent all 10-mers, and most of the sequence represented on it would be redundant.

5.2 LAYING DOWN THE MATRIX

The method described here envisages that the matrix will be produced by synthesising oligonucleotides in the cells of an array by laying down the precursors for the four bases in a predetermined pattern, an example of which is described above. Automatic equipment for applying the precursors has yet to be developed, but there are obvious possibilites; it should not be difficult to adapt a pen plotter or other computer-controlled printing device to the purpose. The smaller the pixel size of the array the better, as complex genomes need very large numbers of cells. However, there are limits to how small these can be made. 100 microns would be a fairly comfortable upper limit, but could probably not be achieved on paper for reasons of texture and diffusion. On a smooth impermeable surface, such as glass, it may be possible to achieve a resolution of around 10 microns, for example by using a laser typesetter to preform a solvent repellant grid, and building the oligonucleotides in the exposed regions. One attractive possibility, which allows adaptation of present techniques of oligonucleotide synthesis, is to sinter microporous glass in microscopic patches onto the surface of a glass plate. Laying down very large number of lines or dots could take a long time, if the printing mechanism were slow. However, a low cost ink-jet printer can print at speeds of about 10,000 spots per second. With this sort of speed, 10^8 spots could be printed in about three hours.

5.3 OLIGONUCLEOTIDE SYNTHESIS

There are several methods of synthesising oligonucleotides. Most methods in current use attach the nucleotides to a solid support of controlled pore size glass (CPG) and are suitable for adaptation to synthesis on a glass surface. Although we know of no description of the direct use of oligonucleotides as hybridisation probes while still attached to the matrix on which they were synthesised, there are reports of the use of oligonucleotides as hybridisation probes on solid supports to which they were attached after synthesis. PCT Application WO 85/01051 describes a method for synthesising oligonucleotides tethered to a CPG column. In an experiment performed by us, CPG was used as the support in an Applied Bio-sytems oligonucleotide synthesiser to synthesise a 13-mer complementary to the left hand cos site of phage lambda. The coupling steps were all close to theoretical yield. The first base was stably attached to the support medium through all the synthesis and deprotection steps by a covalent link.

6. PROBES, HYBRIDISATION AND DETECTION

The yield of oligonucleotides synthesised on microporous glass is about 30 μmol/g. A patch of this material 1 micron thick by 10 microns square would hold ~ 3×10^{-12} μmol, equivalent to about 2 g of human DNA. The hybridisation reaction could therefore be carried out with a very large excess of the bound oligonucleotides over that in the probe. So it should be possible to design a system capable of distinguishing between hybridisation involving single and multiple occurrances of the probe sequence, as yield will be proportional to concentration at all stages in the reaction.

The polynucleotide sequence to be analysed may be of DNA or RNA. To prepare the probe, the polynucleotide may be degraded to form fragments. Preferably it is degraded by a method which is as random as possible, to an average length around the chosen length s of the oligonucleotides on the support, and oligomers of exact length s selected by electrophoresis on a sequencing gel. The probe is then labelled. For example, oligonucleotides of length s may be end labelled. If labelled with ^{32}P, the radioactive yield of any individual s-mer even from total human DNA could be more than 10^4 dpm/mg of total DNA. For detection, only a small fraction of this is needed in a patch 10-100 microns square. This allows hybridisation conditions to be chosen to be close to the Tm of duplexes, which decreases the yield of hybrid and decreases the rate of formation, but increases the discriminating power. Since the bound

oligonucleotide is in excess, signal need not be a problem even working close to equilibrium.

Hybridisation conditions can be chosen to be those known to be suitable in standard procedures used to hybridise to filters, but establishing optimum conditions is important. In particular, temperature needs to be controlled closely, preferably to better than ±0.5°C. Particularly when the chosen length of the oligonucleotide is small, the analysis needs to be able to distinguish between slight differences of rate and/or extent of hybridisation. The equipment may need to be programmed for differences in base composition between different oligonucleotides. In constructing the array, it may be preferable to partition this into sub-matrices with similar base compositions. This may make it easier to define the Tm which may differ slightly according to the base composition.

Autoradiograpy, especially with ^{32}P causes image degradation which may be a limiting factor determining resolution; the limit for silver halide films is around 25 microns. Obviously some direct detection system would be better. Fluorescent probes are envisaged; given the high concentration of the target oligonucleotides, the low sensitivity of fluorescence may not be a problem.

We have considerable experience of scanning autoradiographic images with a digitising scanner. Our present design is capable of resolution down to 25 microns, which could readily be extended down to less than present application, depending on the quality of the hybridisation reaction, and how good it is at distinguishing absence of a sequence from the presence of one or more. Devices for measuring astronomical plates have an accuracy around 1 µm. Scan speeds are such that a matrix of several million cells can be scanned in a few minutes. Software for the analysis of the data is straight-forward, though the large data sets need a fast computer.

Experiments presented below demonstrate the feasibility of the claims.

Commercially available microscope slides (BDH Super Premium 76 x 26 x 1 mm) were used as supports. These were derivatised with a long aliphatic linker that can withstand the conditions used for the deprotection of the aromatic heterocyclic bases, i.e. 30% NH_3 at 55°C for 10 hours. The linker, bearing a hydroxyl group which serves as a starting point for the subsequent oligonucleotide, is synthesised in two steps. The slides are first treated with a 25% solution of 3-glycidoxypropyltriethoxysilane in xylene containing several drops of Hunig's base as a catalyst. The reaction is carried out in a staining jar, fitted with a drying tube, for 20 hours at 90°C. The slides are washed with MeOH, Et_2O and air dried. Then neat hexaethylene glycol and a trace amount of concentrated sulphuric acid are added and the mixture kept at 80°C for 20 hours. The slides are washed with MeOH, Et_2O, air dried and stored desiccated at -20°C until use.

The oligonucleotide synthesis cycle is performed as follows:

The coupling solution is made up fresh for each step by mixing 6 vol. of 0.5M tetrazole in anhydrous acetonitrile with 5 vol. of a 0.2M solution of the required beta-cyanoethylphosphoramidite. Coupling time is three minutes. Oxidation with a 0.1M solution of I_2 in THF/pyridine/H_2O yields a stable phosphotriester bond. Detritylation of the 5' end with 3% trichloroacetic acid in dichloromethane allows further extension of the oligonucleotide chain. There was no capping step since the excess of phosphoramidites used over reactive sites on the slide was large enough to drive the coupling to completion. After the synthesis is completed, the oligonucleotide is deprotected in 30% NH_3 for 10 hours at 55°C. The chemicals used in the coupling step are moisture-sensitive, and this critical step must be performed under anhydrous conditions in a sealed container, as follows. The shape of the patch to be synthesised was cut out of a sheet of silicone rubber (76 x 26 x 0.5 mm) which was sandwiched between a microscope slide, derivatised as described above, and a piece of teflon of the same size and thickness. To this was fitted a short piece of plastic tubing that allowed us to inject and withdraw the coupling solution by syringe and to flush the cavity with argon. The whole assembly was held together by fold-back paper clips. After coupling the set-up was disassembled and the slide put through the subsequent chemical reactions (oxidation with iodine, and detritylation by treatment with TCA) by dipping it into staining jars.

EXAMPLE 1.

As a first example we synthesised the sequences oligo-dT_{10}-oligo-dT_{14} on a slide by gradually decreasing the level of the coupling solution in steps 10 to 14. Thus the 10-mer was synthesised on the upper part of the slide, the 14-mer at the bottom and the 11, 12 and 13-mers were in between. We used 10 pmol oligo-dA_{12}, labelled at the 5' end with ^{32}P by the polynucleotide kinase reaction to a total activity of 1.5 million c.p.m., as a hybridisation probe. Hybridisation was carried out in a perspex (Plexiglas) container made to fit a microscope slide, filled with 1.2 ml of 1M NaCl in TE, 0.1% SDS, for 5 minutes at 20°C. After a short rinse in the same solution without oligonucleotide, we were able to detect more than 2000 c.p.s. with a radiation monitor. An autoradiograph showed that all the counts came from the area where the

oligonucleotide had been synthesised, i.e. there was no non-specific binding to the glass or to the region that had been cerivatised with the linker only. After partial elution in 0.1 M NaCl differential binding to the target is detectable, i.e. less binding to the shorter than the longer oligo-dT. By gradually heating the slide in the wash solution we determined the Tm (mid-point of transition when 50% eluted) to be 33°C. There were no counts detectable after incubation at 39°C. The hybridisation and melting was repeated eight times with no diminution of the signal. The result is reproducible. We estimate that at least 5% of the input counts were taken up by the slide at each cycle.

EXAMPLE 2.

In order to determine whether we would be able to distinguish between matched and mismatched oligonucleotides we synthesised two sequences 3' CCC GCC GCT GGA (cosL) and 3' CCC GCC TCT GGA, which differ by one base at position 7. All bases except the seventh were added in a rectangular patch. At the seventh base, half of the rectangle was exposed in turn to add the two different bases, in two stripes. Hybridisation of cosR probe oligonucleotide (5' GGG CGG CGA CCT) (kinase labelled with ^{32}P to 1.1 million c.p.m., 0.1 M NaCl, TE, 0.1% SDS) was for 5 hours at 32°C. The front of the slide showed 100 c.p.s. after rinsing. Autoradiography showed that annealing occurred only to the part of the slide with the fully complementary oligonucleotide. No signal was detectable on the patch with the mismatched sequence.

EXAMPLE 3.

For a further study of the effects of mismatches or shorter sequences on hybridisation behaviour, we constructed two arrays; one (a) of 24 oligonucleotides and the other (b) of 72 oligonucleotides.

These arrays were set out as shown in Table 1(a) and 1(b). The masks used to lay down these arrays were different from those used in previous experiments. Lengths of silicone rubber tubing (1mm c.d.) were glued with silicone rubber cement to the surface of plain microscope slides, in the form of a "U". Clamping these masks against a derivatised microscope slide produced a cavity into which the coupling solution was introduced through a syringe. In this way only the part of the slide within the cavity came into contact with the phosphoramidite solution. Except in the positions of the mismatched bases, the arrays listed in Table 1 were laid down using a mask which covered most of the width of the slide. Off-setting this mask by 3mm up or down the derivatised slide in subsequent coupling reactions produced the olignucleotides truncated at the 3' or 5' ends.

For the introduction of mismatches a mask was used which covered half (for array (a)) or one third (for array (b)) of the width of the first mask. The bases at positions six and seven were laid down in two or three longitudinal stripes. This led to the synthesis of oligonucleotides differing by one base on each half (array (a)) or third (array (b)) of the slide. In other positions, the sequences differed from the longest sequence by the absence of bases at the ends.

In array (b), there were two columns of sequences between those shown in Table 1(b), in which the sixth and seventh bases were missing in all positions, because the slide was masked in a stripe by the silicone rubber seal. Thus there were a total of 72 different sequences represented on the slide in 90 different positions.

The 19-mer 5' CTC CTG AGG AGA AGT CTG C was used for hybridisation (2 million cpm, 1.2 ml 0.1M NaCl in TE, 0.1% SDS, 20°C).

The washing and elution steps were followed by autoradiography. The slide was kept in the washing solution for 5 min at each elution step and then exposed (45 min, intensified). Elution temperatures were 23, 36, 42, 47, 55 and 60°C respectively.

As indicated in the table, the oligonucleotides showed different melting behaviour. Short oligonucleotides melted before longer ones, and at 55°C, only the perfectly matched 19-mer was stable, all other oligonucleotides had been eluted. Thus the method can differentiate between a 18-mer and a 19-mer which differ only by the absence of one base at the end. Mismatches at the end of the oligonucleotides and at internal sites can all be melted under conditions where the perfect duplex remains.

Thus we are able to use very stringent hybridisation conditions that eliminate annealing to mismatch sequences or to oligonucleotides differing in length by as little as one base. No other method using hybridisation of oligonucleotides bound to the solid supports is so sensitive to the effects of mismatching.

EXAMPLE 4.

To test the application of the invention to diagnosis of inherited diseases, we hybridised the array (a), which carries the oligonucleotide sequences specific for the wild type and the sickle cell mutations of the β-globin gene, with a 110 base pair fragment of DNA amplified from the β-globin gene by means of the polymerase chain reaction (PCR). Total DNA from the blood of a normal individual (1 microgram) was amplified by PCR in the presence of appropriate primer oligonucleotides. The resulting 110 base pair fragment was purified by elecrophoresis through an agarose gel. After elution, a small sample (ca. 10 picogram) was labelled by using α-^{32}P-dCTP (50 microcurie) in a second PCR reaction. This PCR contained only the upstream priming oligonucleotide. After 60 cycles of amplification with an extension time of 9 min the product was removed from precursors by gel filtration. Gel electrophoresis of the radioactive product showed a major band corresponding in length to the 110 base fragment. One quarter of this product (100,000 c.p.m. in 0.9 M NaCl, TE, 0.1% SDS) was hybridised to the array (a). After 2 hours at 30°C ca. 15000 c.p.m. had been taken up. The melting behaviour of the hybrids was followed as described for the 19-mer in example 3, and it was found that the melting behaviour was similar to that of the oligonucleotide. That is to say, the mismatches considerably reduced the melting temperature of the hybrids, and conditions were readily found such that the perfectly matched duplex remained whereas the mismatched duplexes had fully melted.

Thus the invention can be used to analyse long fragments of DNA as well oligonucleotides, and this example shows how it may be used to test nucleic acid sequences for mutations. In particular it shows how it may be applied to the diagnosis of genetic diseases.

EXAMPLE 5.

To test an automated system for laying down the precursors, the cosL oligonucleotide was synthesised with 11 of the 12 bases added in the way described above. For the addition of the seventh base, however, the slide was transferred into an argon filled chamber containing a pen plotter. The pen of the plotter had been replaced by a component, fabricated from Nylon, which had the same shape and dimensions as the pen, but which carried a polytetrafluoroethylene (PTFE) tube, through which chemicals could be delivered to the surface of the glass slide which lay on the bed of the plotter. A microcomputer was used to control the plotter and the syringe pump which delivered the chemicals. The pen, carrying the delivery tube from the syringe, was moved into position above the slide, the pen was lowered and the pump activated to lay down coupling solution. Filling the pen successively with G, T and A phosphoramidite solutions an array of twelve spots was laid down in three groups of four, with three different oligonucleotide sequences. After hybridisation to cosR, as described in Example 2, and autoradiography, signal was seen only over the four spots of perfectly matched oligonucleotides, where the dG had been added.

In conclusion, we have demonstrated the following:

1. It is possible to synthesise oligonucleotides in good yield on a flat glass plate.
2. Multiple sequences can be synthesised on the sample in small spots, at high density, by a simple manual procedure, or automatically using a computer controlled device.
3. Hybridisation to the oligonucleotides on the plate can be carried out by a very simple procedure. Hybridisation is efficient, and hybrids can be detected by a short autoradiographic exposure.
4. Hybridisation is specific. There is no detectable signal on areas of the plate where there are no oligonucleotides. We have tested the effects of mismatched bases, and found that a single mismatched base at any position in oligonucleotides ranging in length from 12-mer to 19-mer reduces the stability of the hybrid sufficiently that the signal can be reduced to a very low level, while retaining significant hybridisation to the perfectly matched hybrid.
5. The oligonucleotides are stably bound to the glass and plates can be used for hybridisation repeatedly.

The invention thus provides a novel way of analysing nucleotide sequences, which should find a wide range of application. We list a number of potential applications below:

Small arrays of oligonucleotides as fingerprinting and mapping tools

Analysis of known mutations including genetic diseases.

Example 4 above shows how the invention may be used to analyse mutations. There are many applications for such a method, including the detection of inherited diseases.

Genomic fingerprinting.

In the same way as mutations which lead to disease can be detected, the method could be used to detect point mutations in any stretch of DNA. Sequences are now available for a number of regions containing the base differences which lead to restriction fragment length polymorphisms (RFLPs). An array of oligonucleotides representing such polymorphisms could be made from pairs of oligonucleotides representing the two allelic restriction sites. Amplification of the sequence containing the RFLP, followed by hybridisation to the plate, would show which alleles were present in the test genome. The number of oligonucleotides that could be analysed in a single analysis could be quite large. Fifty pairs made from selected alleles would be enough to give a fingerprint unique to an individual.

Linkage analysis.

Applying the method described in the last paragraph to a pedigree would pinpoint recombinations. Each pair of spots in the array would give the information that is seen in the track of the RFLP analysis, using gel electrophoresis and blotting, that is now routinely used for linkage studies. It should be possible to analyse many alleles in a single analysis, by hybridisation to an array of allelic pairs of oligonucleotides, greatly simplifying the methods used to find linkage between a DNA polymorphism and phenotypic marker such as a disease gene.

The examples above could be carried out using the method we have developed and confirmed by experiments.

Large arrays of oligonucleotides as sequence reading tools.

We have shown that oligonucleotides can be synthesised in small patches in precisely determined positions by one of two methods: by delivering the precursors through the pen of a pen-plotter, or by masking areas with silicone rubber. It is obvious how a pen plotter could be adapted to synthesise large arrays with a different sequence in each position. For some applications the array should be a predetermined, limited set; for other applications, the array should comprise every sequence of a predetermined length. The masking method can be used for the latter by laying down the precursors in a mask which produces intersecting lines. There are many ways in which this can be done and we give one example for illustration:

1. The first four bases, A, C, G, T, are laid in four broad stripes on a square plate.
2. The second set is laid down in four stripes equal in width to the first, and orthogonal to them. The array is now composed of all sixteen dinucleotides.
3. The third and fourth layers are laid down in four sets of four stripes one quarter the width of the first stripes. Each set of four narrow stripes runs within one of the broader stripes. The array is now composed of all 256 tetranuclectides.
4. The process is repeated, each time laying down two layers with stripes which are one quarter the width of the previous two layers. Each layer added increases the length of the oligonucleotides by one base, and the number of different oligonucleotide sequences by a factor of four.

The dimensions of such arrays are determined by the width of the stripes. The narrowest stripe we have laid is 1mm, but this is clearly not the lowest limit.

There are useful applications for arrays in which part of the sequence is predetermined and part made up of all possible sequences. For example:

Characterising mRNA populations.

Most mRNAs in higher eukaryotes have the sequence AAUAAA close to the 3' end. The array used to analyse mRNAs would have this sequence all over the plate. To analyse a mRNA population it would be hybridised to an array composed of all sequences of the type $N_mAATAAAN_n$. For $m + n = 8$, which should be enough to give a unique oligonucleotide address to most of the several thousand mRNAs that is estimated to be present in a source such as a mammalian cell, the array would be 256 elements square. The 256 x 256 elements would be laid on the AATAAA using the masking method described above. With stripes of around 1mm, the array would be ca. 256mm square.

This analysis would measure the complexity of the mRNA population and could be used as a basis for comparing populations from different cell types. The advantage of this approach is that the differences in the hybridisation pattern would provide the sequence of oligonucleotides that could be used as probes to

isolate all the mRNAs that differed in the populations.

Sequence determination.

To extend the idea to determine unkown sequences, using an array composed of all possible oligonucleotides of a chosen length, requires larger arrays than we have synthesised to date. However, it is possible to scale down the size of spot and scale up the numbers to those required by extending the methods we have developed and tested on small arrays. Our experience shows that the method is much simpler in operation than the gel based methods.

TABLE 1

For Examples 3 and 4 array (a) was set out as follows:

20	GAG GAC TCC TCT ACG	20	GAG GAC aCC TCT ACG
36	GAG GAC TCC TCT GAC G	20	GAC GAC aCC TCT GAC G
36	GAG GAC TCC TCT AGA CG	20	GAC GAC aCC TCT AGA CG
47	GAG GAC TCC TCT CAG ACG	36	GAG GAC aCC TCT CAG ACG
60	GAG GAC TCC TCT TCA GAC G	47	GAG GAC aCC TCT TCA GAC G
56	.AG GAC TCC TCT TCA GAC G	42	.AG GAC aCC TCT TCA GAC G
56	..G GAC TCC TCT TCA GAC G	42	..G GAC aCC TCT TCA GAC G
47	... GAC TCC TCT TCA GAC G	42	... GAC aCC TCT TCA GAC G
42	... .AC TCC TCT TCA GAC G	36	... .AC aCC TCT TCA GAC G
36	... ..C TCC TCT TCA GAC G	36	... ..C aCC TCT TCA GAC G
36	... ... TCC TCT TCA GAC G	36	... ... aCC TCT TCA GAC G
36	... ... .CC TCT TCA GAC G	36	... ... .CC TCT TCA GAC G

For example 3 array (b) was set out as follows:

20 GAG GAt TC	20 GAG GAC TC	20 GAG GAC aC
20 GAG GAt TCC	20 GAG GAC TCC	20 GAG GAC aCC
20 GAG GAt TCC T	20 GAG GAC TCC T	20 GAG GAC aCC T
20 GAG GAt TCC TC	20 GAG GAC TCC TC	20 GAG GAC aCC TC
20 GAG GAt TCC TCT	20 GAG GAC TCC TCT	20 GAG GAC aCC TCT
20 GAG GAt TCC TCT T	20 GAG GAC TCC TCT T	20 GAG GAC aCC TCT T
20 GAG GAt TCC TCT TC	20 GAG GAC TCC TCT TC	20 GAG GAC aCC TCT TC
20 GAG GAt TCC TCT TCA	20 GAG GAC TCC TCT TCA	20 GAG GAC aCC TCT TCA
32 GAG GAt TCC TCT TCA G	42 GAG GAC TCC TCT TCA G	20 GAG GAC aCC TCT TCA G
32 GAG GAt TCC TCT TCA GA	47 GAG GAC TCC TCT TCA GA	32 GAG GAC aCC TCT TCA GA
42 GAG GAt TCC TCT TCA GAC	52 GAG GAC TCC TCT TCA GAC	42 GAG GAC aCC TCT TCA GAC
52 GAG GAt TCC TCT TCA GAC G	60 GAG GAC TCC TCT TCA GAC G	52 GAG GAC aCC TCT TCA GAC G
42 .AG GAt TCC TCT TCA GAC G	52 .AG GAC TCC TCT TCA GAC G	42 .AG GAC aCC TCT TCA GAC G
42 ..G GAt TCC TCT TCA GAC G	52 ..G GAC TCC TCT TCA GAC G	42 ..G GAC aCC TCT TCA GAC G
37 ... GAt TCC TCT TCA GAC G	47 ... GAC TCC TCT TCA GAC G	37 ... GAC aCC TCT TCA GAC G
32 ... .At TCC TCT TCA GAC G	42 ... .AC TCC TCT TCA GAC G	32 ... .AC aCC TCT TCA GAC G
32 ... ..t TCC TCT TCA GAC G	42 ... ..C TCC TCT TCA GAC G	32 ... ..C aCC TCT TCA GAC G
32 ... ... TCC TCT TCA GAC G	32 ... ... TCC TCT TCA GAC G	32 ... ... aCC TCT TCA GAC G

Between the three columns of array (b) listed above, were two columns, in which bases 6 and 7 from the left were missing in every line. These sequences all melted at 20 or 32 degrees. (a,t) mismatch base (.) missing base.

Claims

1. A method of analyzing a polynucleotide sequence, by the use of a support to the surface of which is attached an array of the whole or a chosen part of a complete set of oligonucleotides of chosen

lengths, the different oligonucleotides occupying separate cells of the array, which method comprises labelling the polynucleotide sequence or fragments thereof, applying the polynucleotide sequence or fragments thereof under hybridisation conditions to the array, and observing the location of the label on the surface associated with particular members of the set of oligonucleotides.

2. A method according to claim 1, applied to the study of differences between polynucleotide sequences, wherein the array is of the whole or a chosen part of the complete set of oligonucleotides of chosen lengths comprising the polynucleotide sequences.

3. A method as claimed in claim 2, wherein the array comprises one or more pairs of oligonucleotides of chosen lengths.

4. A method as claimed in claim 3, wherein the array comprises one or more pairs of oligonucleotides representing normal and mutant versions of a point mutation being studied.

5. A method according to any one of claims 1 to 4, wherein the polynucleotide sequence is randomly degraded to form a mixture of oligomers of a chosen length, the mixture being thereafter labelled to form labelled material which is applied to the array.

6. A method as claimed in claim 5, wherein the oligomers are labelled with ^{32}P.

7. A method as claimed in any one of claims 1 to 6, wherein the chosen length of the oligonucleotides is from 8 to 20 nucleotides.

8. A method as claimed in any one of claims 1 to 7, wherein each oligonucleotide is bound to the support by a covalent link through a terminal nucleotide.

9. A method as claimed in any one of claims 1 to 8, wherein the oligonucleotide at each cell has a defined sequence.

10. Apparatus suitable for analysing a polynucleotide sequence by the method of any one of claims 1 to 9 comprising a support and attached to a surface thereof an array of the whole or a chosen part of a complete set of oligonucleotides of chosen lengths, the different oligonucleotides occupying separate cells of the array and being capable of taking part in hybridisation reactions.

11. Apparatus as claimed in claim 10 for studying differences between polynucleotide sequences, wherein the array is of the whole or a chosen part of a complete set of oligonucleotides of chosen lengths comprising the polynucleotide sequences.

12. Apparatus as claimed in claim 11, wherein the array comprises one or more pairs of oligonucleotides of chosen lengths.

13. Apparatus as claimed in claim 12, wherein the array comprises one or more pairs of oligonucleotides representing normal and mutant versions of a point mutation to be studied.

14. Apparatus for determining the sequence of a polynucleotide comprising a support having attached to a surface thereof an array of different oligonucleotides with defined sequences, the oligonucleotides occupying cells of the array and being attached to the surface, wherein the defined sequence of an oligonucleotide of one cell of the array is different than the defined sequence of an oligonucleotide of another cell of the array.

15. Apparatus for analysing a polynucleotide, the apparatus comprising a support segregated into at least two defined cells, each cell having attached thereto oligonucleotides with known sequence, where the sequence of the oligonucleotides of a first cell is different than the sequence of the oligonucleotides of a different cell.

16. Apparatus as claimed in any one of claims 10 to 15, wherein the chosen length of the oligonucleotides is from 8 to 20 nucleotides.

17. Apparatus as claimed in any one of claims 10 to 16, wherein the surface of the support to which the oligonucleotides are attached is of glass.

18. Apparatus as claimed in any one of claims 10 to 17, wherein each oligonucleotide is bound to the support by a covalent link through a terminal nucleotide.

19. A method for generating, for the apparatus of claim 14, an array of oligonucleotides of chosen lengths within discrete cells of a support material comprising the steps of
a) segregating a support material into discrete cell locations;
b) coupling a nucleotide to a first set of cell locations;
c) coupling a nucleotide to a second set of cell locations;
d) coupling a nucleotide to a third set of cell locations;
e) and continuing the sequence of coupling steps until the desired array has been generated, the coupling being effected at each location either to the surface of the support or to a nucleotide coupled in a previous step at that location.

20. The method of claim 19 wherein a microcomputer controlled plotter delivers the nucleotides to said sets of cell locations.

21. The method of claim 19 or claim 20 wherein the size of each discrete cell is between 10 and 100 μm.

22. The method of any one of claims 19 to 21 further comprising the use of means for coupling said nucleotides to a particular set of discrete cell locations to the exclusion of other discrete cell locations.

23. The method of claim 22 wherein said means is a mask.

Patentansprüche

1. Verfahren zur Analyse einer Polynukleotid-Sequenz, bei dem ein Träger verwendet wird, an dessen Oberfläche eine Anordnung des gesamten oder eines ausgewählten Teils einer vollständigen Gruppe von Oligonukleotiden ausgewählter Längen gebunden ist, wobei die verschiedenen Oligonukleotide separate Zellen der Anordnung besetzen, wobei bei dem Verfahren die Polynukleotid-Sequenz oder Fragmente davon markiert werden, die Polynukleotid-Sequenz oder Fragmente davon unter Hybridisierungsbedingungen auf die Anordnung aufgebracht werden und die Lokalisierung der Markierung auf der Oberfläche, die mit bestimmten Bestandteilen der Gruppe von Oligonukleotiden verbunden ist, beobachtet wird.

2. Verfahren nach Anspruch 1, verwendet zur Untersuchung von Unterschieden zwischen Polynukleotid-Sequenzen, bei dem die Anordnung aus dem gesamten oder einem ausgewählten Teil der vollständigen Gruppe von Oligonukleotiden ausgewählter Längen besteht, welche die Polynukleotid-Sequenzen aufweist.

3. Verfahren nach Anspruch 2, wobei die Anordnung ein oder mehrere Paare von Oligonukleotiden ausgewählter Längen aufweist.

4. Verfahren nach Anspruch 3, wobei die Anordnung ein oder mehrere Paare von Oligonukleotiden aufweist, die Normal- und Mutantenversionen einer untersuchten Punktmutation darstellen.

5. Verfahren nach einem der Ansprüche 1 bis 4, wobei die Polynukleotid-Sequenz statistisch zur Bildung eines Gemisches von Oligomeren ausgewählter Längen abgebaut wird, wobei das Gemisch danach zur Bildung von markiertem Material, das auf die Anordung aufgebracht wird, markiert wird.

6. Verfahren nach Anspruch 5, wobei die Oligomeren mit ^{32}P markiert werden.

7. Verfahren nach einem der Ansprüche 1 bis 6, wobei die ausgewählte Länge der Oligonukleotide 8 bis 20 Nukleotide beträgt.

8. Verfahren nach einem der Ansprüche 1 bis 7, wobei jedes Oligonukleotid mit einer kovalenten Bindung durch ein terminales Nukleotid an den Träger gebunden ist.

9. Verfahren nach einem der Ansprüche 1 bis 8, wobei das Oligonukleotid in jeder Zelle eine definierte Sequenz besitzt.

10. Vorrichtung, welche zur Analyse einer Polynukleotid-Sequenz mit dem Verfahren nach einem der Ansprüche 1 bis 9 geeignet ist, die einen Träger und an eine Oberfläche davon gebunden eine Anordnung des gesamten oder eines ausgewählten Teils einer vollständigen Gruppe von Oligonukleotiden ausgewählter Längen aufweist, wobei die verschiedenen Oligonukleotide separate Zellen der Anordnung besetzen und an Hybridisierungsreaktionen teilnehmen können.

11. Vorrichtung nach Anspruch 10 zur Untersuchung von Unterschieden zwischen Polynukleotid-Sequenzen, bei dem die Anordnung aus dem gesamten oder einem ausgewählten Teil einer vollständigen Gruppe von Oligonukleotiden ausgewählter Längen besteht, welche die Polynukleotid-Sequenzen aufweist.

12. Vorrichtung nach Anspruch 11, wobei die Anordnung ein oder mehrere Paare von Oligonukleotiden ausgewählter Längen aufweist.

13. Vorrichtung nach Anspruch 12, wobei die Anordnung ein oder mehrere Paare von Oligonukleotiden aufweist, die Normal- und Mutantenversionen einer zu untersuchenden Punktmutation darstellen.

14. Vorrichtung zur Bestimmung der Sequenz eines Polynukleotids, die einen Träger aufweist, an dessen eine Oberfläche eine Anordnung von verschiedenen Oligonukleotiden mit definierten Sequenzen gebunden ist, wobei die Oligonukleotide Zellen der Anordnung besetzen und an die Oberfläche gebunden sind, wobei die definierte Sequenz eines Oligonukleotids einer Zelle der Anordnung sich von der definierten Sequenz eines Oligonukleotids einer anderen Zelle der Anordnung unterscheidet.

15. Vorrichtung zur Analyse eines Polynukleotids, wobei die Vorrichtung einen Träger aufweist, der in mindestens zwei definierte Zellen aufgeteilt ist, wobei an jede Zelle Oligonukleotide mit bekannter Sequenz gebunden sind, wobei die Sequenz der Oligonukleotide einer ersten Zelle sich von der Sequenz der Oligonukleotide einer anderen Zelle unterscheidet.

16. Vorrichtung nach einem der Ansprüche 10 bis 15, wobei die ausgewählte Länge der Oligonukleotide 8 bis 20 Nukleotide beträgt.

17. Vorrichtung nach einem der Ansprüche 10 bis 16, wobei die Oberfläche des Trägers, an den die Oligonukleotide gebunden sind, aus Glas ist.

18. Vorrichtung nach einem der Ansprüche 10 bis 17, wobei jedes Oligonukleotid mit einer kovalenten Bindung durch ein terminales Nukleotid an den Träger gebunden ist.

19. Verfahren zur Erzeugung, nämlich für die Vorrichtung nach Anspruch 14, einer Anordnung von Oligonukleotiden ausgewählter Längen in diskreten Zellen eines Trägermaterials, bei dem
a) ein Trägermaterial in diskrete Zell-Lokalisierungen aufgeteilt wird;
b) ein Nukleotid an eine erste Gruppe von Zell-Lokalisierungen gekuppelt wird;
c) ein Nukleotid an eine zweite Gruppe von Zell-Lokalisierungen gekuppelt wird;
d) ein Nukleotid an eine dritte Gruppe von Zell-Lokalisierungen gekuppelt wird; und
e) die Sequenz der Kupplungsschritte fortgesetzt wird, bis die gewünschte Anordnung erzeugt ist,
wobei die Kupplung an jeder Lokalisierung entweder an die Oberfläche des Trägers oder an ein in einem vorherigen Schritt an diese Lokalisierung gekuppeltes Nukleotid erfolgt.

20. Verfahren nach Anspruch 19, wobei ein Mikrocomputer-gesteuerter Plotter die Nukleotide an die Gruppen von Zell-Lokalisierungen abgibt.

21. Verfahren nach Anspruch 19 oder Anspruch 20, wobei die Größe jeder diskreten Zelle zwischen 10 und 100 µm liegt.

22. Verfahren nach einem der Ansprüche 19 bis 21, bei dem außerdem eine Einrichtung zur Kupplung der Nukleotide an eine bestimmte Gruppe von diskreten Zell-Lokalisierungen unter Ausschluß von anderen diskreten Zell-Lokalisierungen verwendet wird.

23. Verfahren nach Anspruch 22, wobei es sich bei der Einrichtung um eine Maske handelt.

Revendications

1. Procédé d'analyse d'une séquence de polynucléotide par utilisation d'un support à la surface duquel est attaché une série de la totalité ou d'une partie d'un ensemble complet d'oligonucléotides de longueurs choisies, les différents oligonucléotides occupant des cellules séparées de la série, dans lequel on marque la séquence de polynucléotide ou des fragments de cette séquence, on applique la séquence de polynucléotide ou ses fragments dans des conditions d'hybridation à la série, et on observe la localisation de la marque à la surface associée à des membres particuliers de l'ensemble d'oligonucléotides.

2. Procédé selon la revendication 1, appliqué à l'étude des différences entre les séquences de polynucléotides, où la série est celle de la totalité ou d'une partie choisie de l'ensemble complet d'oligonucléotides de longueurs choisies constituant les séquences de polynucléotides.

3. Procédé selon la revendication 2, dans lequel la série comprend une ou plusieurs paires d'oligonucléotides de longueurs choisies.

4. Procédé selon la revendication 3, dans lequel la série comprend une ou plusieurs paires d'oligonucléotides représentant les versions normale et mutante d'une mutation ponctuelle que l'on étudie.

5. Procédé selon l'une quelconque des revendications 1 à 4, dans lequel la séquence de polynucléotide est dégradée de façon aléatoire pour former un mélange d'oligomères d'une longueur choisie, le mélange étant ensuite marqué pour former une matière marquée qui est appliquée à la série.

6. Procédé selon la revendication 5, dans lequel les oligomères sont marqués avec ^{32}P.

7. Procédé selon l'une quelconque des revendications 1 à 6, dans lequel la longueur choisie des oligonucléotides est de 8 à 20 nucléotides.

8. Procédé selon l'une quelconque des revendications 1 à 7, dans lequel chaque oligonucléotide est lié au support par une liaison de covalence par l'intermédiaire d'un nucléotide terminal.

9. Procédé selon l'une quelconque des revendications 1 à 8, dans lequel l'oligonucléotide à chaque cellule a une séquence définie.

10. Appareil approprié pour analyser une séquence de polynucléotide par le procédé de l'une quelconque des revendications 1 à 9 comprenant un support et, attachée à une de ses surfaces, une série de la totalité ou d'une partie choisie d'un ensemble complet d'oligonucléotides de longueurs choisies, les différents oligonucléotides occupant des cellules séparées de la série et étant capables de prendre part à des réactions d'hybridation.

11. Appareil selon la revendication 10 pour étudier les différences entre les séquences de polynucléotides, où la série est celle de la totalité ou d'une partie choisie d'un ensemble complet d'oligonucléotides de longueurs choisies constituant les séquences de polynucléotides.

12. Appareil selon la revendication 11, dans lequel la série comprend une ou plusieurs paires d'oligonucléotides de longueurs choisies.

13. Appareil selon la revendication 12, dans lequel la série comprend une ou plusieurs paires d'oligonucléotides représentant les versions normale et mutante d'une mutation ponctuelle à étudier.

14. Appareil pour déterminer la séquence d'un polynucléotide comprenant un support à une surface duquel est attachée une série d'oligonucléotides différents ayant des séquences définies, les oligonucléotides occupant des cellules de la série et étant attachés à la surface, où la séquence définie d'un oligonucléotide d'une cellule de la série est différente de la séquence définie d'un oligonucléotide d'une autre cellule de la série.

15. Appareil pour analyser un polynucléotide, l'appareil comprenant un support séparé en au moins deux cellules définies, avec fixation sur chaque cellule d'oligonucléotides de séquences connues, où la séquence des oligonucléotides d'une première cellule est diférente de la séquence des oligonucléotides d'une cellule différente.

16. Appareil selon l'une quelconque des revendications 10 à 15, dans lequel la longueur choisie des oligonucléotides est de 8 à 20 nucléotides.

17. Appareil selon l'une quelconque des revendications 10 à 16, dans lequel la surface du support auquel les oligonucléotides sont attachés est de verre.

18. Appareil selon l'une quelconque des revendications 10 à 17, dans lequel chaque oligonucléotide est lié au support par une liaison de covalence par l'intermédiaire d'un nucléotide terminal.

19. Procédé pour générer, pour l'appareil de la revendication 14, une série d'oligonucléotides de longueurs choisies dans des cellules discrètes d'une matière support, comprenant les étapes de
a) séparation d'une matière support en localisations de cellules discrètes;
b) couplage d'un nucléotide à un premier ensemble de localisations de cellules;
c) couplage d'un nucléotide à un second ensemble de localisations de cellules;
d) couplage d'un nucléotide à un troisième ensemble de localisations de cellules;
e) et poursuite de la séquence d'étapes de couplage jusqu'à ce que la série désirée ait été générée, le couplage étant effectué à chaque localisation soit à la surface du support, soit à un nucléotide couplé dans une étape précédente à cette localisation.

20. Procédé de la revendication 19 dans lequel un appareil de restitution commandé par micro-ordinateur délivre les nucléotides auxdits ensembles de localisations de cellules.

21. Procédé de la revendication 19 ou de la revendication 20 dans lequel la taille de chaque cellule discrète est comprise entre 10 et 100 µm.

22. Procédé de l'une quelconque des revendications 19 à 21 comprenantt en outre l'utilisation de moyens pour coupler lesdits nucléotides à un ensemble particulier de localisations de cellules discrètes à l'exclusion d'autres localisations de cellules discrètes.

23. Procédé de la revendication 22 dans lequel ledit moyen est un masque.

BESCHWERDEKAMMERN DES EUROPÄISCHEN PATENTAMTS | BOARDS OF APPEAL OF THE EUROPEAN PATENT OFFICE | CHAMBRES DE RECOURS DE L'OFFICE EUROPEEN DES BREVETS

Internal distribution code:
(A) [ ] Publication in OJ
(B) [ ] To Chairmen and Members
(C) [x] To Chairmen
(D) [ ] No distribution

DECISION
of 12 October 2005

Case Number: T 0378/02 - 3.3.04

Application Number: 89905449.8

Publication Number: 0373203

IPC: C12Q 1/68

Language of the proceedings: EN

Title of invention:
Method and apparatus for analysing polynucleotide sequences

Patentee:
OXFORD GENE TECHNOLOGY LIMITED

Opponents:
01: Hyseq Inc.
02: Nanogen Inc.
03: F.HOFFMANN-LA ROCHE & CO.
04: Abbott Laboratories
05: Multilyte Ltd.
06: Vysis Corp.

Headword:
Analysis of polynucleotide sequences/OXFORD GENE TECHNOLOGY

Relevant legal provisions:
EPC Art. 54, 56, 83, 84, 123(2)(3)

Keyword:
"Novelty, inventive step, sufficiency of disclosure, clarity (yes), added subject-matter, extension of scope of protection (no)"

Decisions cited:
T 0019/90, T 0409/91, T 0694/92, T 0412/93, T 0860/93, T 0639/95, T 0860/95, T 0994/95, T 0188/97, T 0636/97, T 0649/97, T 0728/98, T 1041/98, T 0193/01

Catchword:
-

Case Number: T 0378/02 - 3.3.04

DECISION
of the Technical Board of Appeal 3.3.04
of 12 October 2005

Appellant I: (Opponent 02)
Nanogen Inc.
10398 Pacific Center Court
San Diego, CA92121 (US)

Representative:
Taormino, Joseph Paul
Hoffmann - Eitle
Patent- und Rechtsanwälte
Arabellastrasse 4
D-81925 München (DE)

Appellant II: (Opponent 03)
F. HOFFMANN-LA ROCHE & CO
Aktiengesellschaft
Grenzacherstrasse 124
CH-4002 Basel (CH)

Representative:
Jaenichen, Hans-Rainer
Vossius & Partner
Postfach 86 07 67
D-81634 München (DE)

Appellant III: (Opponent 05)
Multilyte Ltd.
c/o MJ Ventham & Co.
2nd. floor Kingsbourne House
229-231 High Holborn
LONDON WC1V 7DA (GB)

Representative:
Armitage, Ian Michael
Mewburn Ellis LLP
York House
23 Kingsway
London WC2B 6HP (GB)

Appellant IV: (Opponent 06)
Vysis Corp.
3100 Woodcreek Drive
Downers Grove IL 80515 (US)

Representative:
Ritter, Stephen David
Mathys & Squire
120 Holborn
London EC1N 2SQ (GB)

Respondent: (Patent Proprietor)	OXFORD GENE TECHNOLOGY LIMITED 12 School Road Kidlington, Oxford OX5 2HB (GB)
Representative:	Hallybone, Huw George Carpmaels and Ransford, 43 Bloomsbury Square London WC1A 2RA (GB)
Other Parties: (Opponent 01)	Hyseq Inc. 670 Almanor Ave. Sunnyvale CA 94086 (US)
Representative:	Brown, John D. FORRESTER & BOEHMERT Pettenkoferstrasse 20-22 D-80336 München (DE)
(Opponent 04)	Abbott Laboratories One Abbott Park Road Abbott Park, Illinois 60064 (US)
Representative:	Ritter, Stephen David Mathys & Squire 120 Holborn London EC1N 2SQ (GB)
Decision under appeal:	**Interlocutory decision of the Opposition Division of the European Patent Office posted 26 February 2002 concerning maintenance of the European patent No. 0373203 in amended form.**

Composition of the Board:

Chairman: M. Wieser
Members: G. Alt
G. Weiss

Summary of Facts and Submissions

I. European patent No. 0 373 203 with the title "Method and apparatus for analysing polynucleotide sequences" was granted with twenty-three claims on the basis of European patent application 89 905 449.8 which was derived from International application WO 89/10977.

II. The patent had been opposed by seven parties (opponents 01 to 07) under Article 100(a) EPC for not being an invention in the sense of Article 52(2) EPC, for lack of novelty (Article 54 EPC) and lack of inventive step (Article 56 EPC), under Article 100(b) EPC on the ground of lack of sufficient disclosure (Article 83 EPC) and under Article 100(c) EPC on the ground of added subject-matter.

Opponent 07 withdrew his opposition during the opposition procedure and thus ceased to be a party to the procedure.

III. The opposition division had decided that the claims of the main request before them violated the requirements of Article 123(2) EPC, but that the claims of the first auxiliary request met all requirements of the EPC.

IV. Appeals were lodged by opponent 02 (appellant I), opponent 03 (appellant II), opponent 05 (appellant III) and opponent 06 (appellant IV).

Opponents 01 and 04 are parties to the proceedings as of right according to Article 107 EPC.

V. The patent proprietor, who initially also lodged an appeal against the decision of the opposition division, withdrew this appeal with letter of 10 October 2005 and is therefore respondent in the present appeal procedure.

VI. The board expressed its preliminary opinion in a communication dated 14 July 2005. Oral proceedings were held on 11 and 12 October 2005 in the absence of appellants II and IV and of the other parties, opponents 01 and 04. At these proceedings the respondent filed a new main request consisting of claims 1 to 22.

Independent claims 1, 10, 14, 15 and 18 thereof read as follows:

"1. A method of analysing a polynucleotide sequence by the use of a glass support, to a smooth impermeable surface of which is attached an array of the whole or a chosen part of a complete set of oligonucleotides of chosen lengths, the different oligonucleotides being attached through a covalent link and occupying separate cells of the array, which method comprises labelling the polynucleotide sequence or fragments thereof, applying the polynucleotide sequence or fragments thereof under hybridisation conditions to the array, and observing the location of the label on the surface associated with particular members of the set of oligonucleotides.

10. Apparatus suitable for analysing a polynucleotide sequence by the method of any one of claims 1 to 9 comprising a glass support and attached to a smooth impermeable surface thereof an array of the whole or a

chosen part of a complete set of oligonucleotides of chosen lengths, the different oligonucleotides being attached through a covalent link, occupying separate cells of the array, and being capable of taking part in hybridisation reactions.

14. Apparatus for determining the sequence of a polynucleotide comprising a glass support having attached to a smooth impermeable surface thereof an array of different oligonucleotides with defined sequences, the oligonucleotides occupying cells of the array and being attached through a covalent link to the surface and being capable of taking part in hybridisation reactions, wherein the defined sequence of an oligonucleotide of one cell of the array is different than the defined sequence of an oligonucleotide of another cell of the array.

15. Apparatus for analysing a polynucleotide, the apparatus comprising a glass support segregated into at least two defined cells, each cell having attached to a smooth impermeable surface thereof, through a covalent link, oligonucleotides with known sequence, capable of taking part in hybridisation reactions, where the sequence of the oligonucleotides of a first cell is different than the sequence of the oligonucleotides of a different cell.

18. A method for generating, for the apparatus of claim 14, an array of oligonucleotides of chosen lengths within discrete cells of a glass support material having a smooth impermeable surface comprising the steps of

a) segregating the smooth impermeable surface of the support material into discrete cell locations;
b) coupling a nucleotide to a first set of cell locations;
c) coupling a nucleotide to a second set of cell locations;
d) coupling a nucleotide to a third set of cell locations;
e) and continuing the sequence of coupling steps until the desired array has been generated,

the coupling being effected at each location either to the surface of the support or to a nucleotide coupled in a previous step at that location."

VII. Appellants I to IV requested that the decision under appeal be set aside and that the patent be revoked.

The respondent requested that the decision under appeal be set aside and that the patent be maintained on the basis of claims 1 to 22 of the main request filed at the oral proceedings on 11 October 2005.

VIII. The following documents are referred to in this decision:

OD1:	EP-A-0 235 726
OD2:	EP-A-0 171 150
OD6:	Cell, col. 12, 1977, pages 23-36, Dunn, A. and Hassell, J.
OD13:	WO 85/01051
OD15:	Nucleic Acids Research, vol. 11, no. 13, 1983, pages 4365-4377, Frank, R. et al.
OD18:	WO 88/01302

OD19: WO 86/03782

OD20: US 4,395,486

OD32: WO 93/22480

OD33: Genomics, vol. 13, 1992, pages 1008-1017, Southern, E.M. et al.

OD38: Nucleic Acids Research, vol. 15, no. 13, 1987, pages 5373-5390, Gingeras, T.R. et al.

OD41: US 4,704,353

OD45: WO 84/03151

OD50: EP-A-0 130 523

OD61: Nucleic Acids Research, vol. 15, no. 13, July 1987, pages 5353-5371, Ghosh, S. S. and Musso, G.F.

OD87: US 4,591,570

AD95: Gelb, L.V. and Gubbins, K.E. "Characterization of porous glasses by adsorption: Models, simulations and data inversion" referring to Nature, vol. 206, 1965, pages 693-696, Haller, W. (reference number [4])

AD96: Nucleic Acids Research, vol. 15, no. 7, 1987,pages 2891-2909, Kremsky, J.N. et al.

AD97: Nucleic Acids Research, vol. 15, no. 7, 1987, pages 2911-2916, Wolf, S.F. et al.

AD100: Nature Biotechnology, vol. 15, December 1997, pages 1359-1367

AD101: US 4,000,252

AD102: US 4,145,406

AD103: US 4,205,952

AD104: US 4,254,082

AD106: US 4,323,647

AD107: US 4,442,204

Lecture by Prof Ekins given on 11 April 1988 and post-published paper corresponding to said lecture

IX. The arguments of appellants I to IV as far as they are relevant for the present decision may be summarized as follows:

Clarity

The term "...a glass support, to a smooth impermeable surface of which is attached..." not only encompassed the covalent attachment of the oligonucleotides directly to the glass surface, but also the indirect attachment through a second layer situated on top of the glass surface. This second layer was defined only by the two relative terms "smooth" and "impermeable", which made it impossible to exactly define the nature of this layer.

Even if the terms "smooth" and "impermeable" were considered to define the direct surface of the glass support, this definition was ambiguous because of the existence of different kinds of glass with different degrees of smoothness.

Novelty

Document OD2 described a method for screening by hybridisation of a plurality of unknown nucleic acids covalently attached to a solid support. Several materials were disclosed as support materials including, inter alia, "glass (e.g. solid, fibre etc)" (page 6). This definition implicitly included glass with a smooth and impermeable surface. Since the document also disclosed all the other features of the subject-matter of claims 1, 14 and 15, it destroyed the novelty.

Sufficiency of disclosure

The claimed method embraced high density arrays. However, the specification failed to enable the production of such arrays because, as stated in the patent in suit, the automatic equipment for achieving them had not yet been produced at the priority date and the level of time and effort needed for the adaptation of existing equipment according to the suggestions in the patent in suit amounted to an undue burden. According to decisions T 994/95, T 188/97, T 412/93 and T 639/95 sufficiency of disclosure should be denied under such circumstances.

Statements in post-published documents OD32 and OD33, saying that large-scale nucleic acid sequence analysis was not achievable due to the lack of automated equipment and that even with automated methods synthesis of a large number of oligonucleotides was not easy, corroborated the view that the claimed invention had not been sufficiently disclosed at the priority date of the patent.

Document AD100, published in 1997, dealt with a high-density array for monitoring the expression of the yeast genome and was therefore exemplary for modern array techniques. However, some aspects of the method disclosed therein were not implemented in the method of the patent in suit. This discrepancy demonstrated that the invention was not sufficiently disclosed.

Inventive step

Either the lecture of Prof. Ekins or one of documents OD1 or OD2 was the closest prior art document.

In his lecture Prof. Ekins reported about a multi-analyte microspot immunoassay system allowing the determination of a large number of different proteins in a sample and using glass as a support. Starting from this prior art it would have been obvious to the skilled person that the same multispot format could be used for nucleic acid hybridisation assays since it was well-known that antibodies and nucleic acids were alternative types of binding agents that could be used interchangeably in assays as suggested in documents AD101 to AD108.

Documents OD1 or OD2 both related to nucleic acid assays in array format. The problem to be solved in the light of their disclosure was the provision of a further, alternative assay. The solution to this problem was the use of a support made of glass with a smooth and impermeable surface.

This solution was obvious in view of either of documents OD13 or OD61 disclosing methods to covalently immobilise DNA on controlled pore glass. A skilled person would have realized that these techniques were also suitable for coupling of nucleic acid to glass with a smooth impermeable surface and would therefore have replaced the support material of documents OD1 and OD2 by that type of material.

Moreover, the aspect of attaching a receptor and an analyte in a regular pattern on a surface was not confined to the situation where both receptor and analyte were nucleic acids. Therefore, it was legitimate to look at disclosures relating to the attachment of a different substance to a surface, for example at documents OD45 or the related US patent OD87 dealing with immunoassays and mentioning specifically the use of a "flat, planar surface such as glass or a plastic coverslip". Documents such as AD101 to AD108 relating generally to binding assays confirmed that nucleic acid and antibody antigen binding were equivalent and that the same support material could be used.

Finally, document OD41, relating to photoresponsive redox detection, disclosed that oligonucleotides could be bound to glass surfaces.

X. The respondent's arguments as far as they are relevant for the present decision may be summarized as follows:

Clarity

Relative terms in claims are allowable if they are clear in view of the complete specification, see for example decisions T 860/93, T 860/95, T 649/97, T 1041/98 and T 193/01. This was the case here, because it was clear from the disclosure in the patent as a whole that the definition "a glass support, to a smooth and impermeable surface of which" meant ordinary glass like window glass or glass of which microscope slides were made.

Moreover, a second layer on the top of the glass surface was nowhere disclosed in the patent in suit.

Novelty

Document OD2 did not disclose smooth impermeable glass as a support because the type of glass referred to in document OD2 had to be flexible. This was not a property of the type of glass defined in claim 1 of the patent in suit.

Sufficiency of disclosure

The appellants did not provide evidence substantiating difficulties when repeating examples of the patent or proving errors in its technical details. Without such evidence an objection under Article 83 EPC could not be successful in view of decision T 19/90.

It may be true that document AD100 disclosed process parameters which were different from those in the patent in suit. However, showing that something worked in a different way was not a proof that it did not also work the other way.

Inventive step

The lecture by Prof. Ekins could not be considered to be the closest prior art document because it did not relate to nucleic acid hybridisation. Rather, the closest prior art was represented by document OD1. The patent in suit was distinguished therefrom by the use of a support made of glass with a smooth impermeable surface for covalent attachment of nucleic acids.

The problem to be solved was the provision of an improved apparatus and method for parallel analysis of nucleic acid by hybridisation.

Up to the time when the invention was made, porous supports were considered to be necessary since they could bind more nucleic acid. Neither was there a hint in the closest prior art, document OD1, to use a different support material, nor could the skilled person, trying to solve the posed problem, arrive at the claimed solution in an obvious way by combining the teaching in OD1 with any other prior art document on file because none of them pointed to the covalent attachment of nucleic acid to glass with a smooth and impermeable surface.

Reasons for the Decision

Amendments and extension of scope

1. The claims have the following basis in the application documents as originally filed:

Claim 1 differs from claim 8 as originally filed by the term "glass" in front of the term "support". This amendment and the corresponding amendments in claims 10, 14, 15 and 18 are based on page 11, lines 24 to 25 and claim 6 as originally filed.

Claims 2 to 7 are based on claims 9 to 14 and claim 8 is based on claim 7 as originally filed.

Support for claim 9 comes from the application documents as originally filed as a whole because the gist of the method disclosed therein is that the bound known sequence is tested with an unknown sequence. The presence of oligonucleotides in cells is disclosed on page 11: "The method described here envisages that the matrix will be produced by synthesising oligonucleotides in the **cells** of an array...".

Claim 10 is a combination of claims 1 and 8 as originally filed.

Claims 11 to 13 correspond to claims 2 to 4 as originally filed.

Claim 14 is based on claim 1 and point 4 of the description as originally filed. The expression "capable of taking part in hybridization reactions" which is likewise introduced in claim 15, is found in claim 1 as originally filed.

Claim 15 relies on Example 2 disclosing the attachment of two oligonucleotides with different sequences to the surface of the support and claim 3 referring to an array comprising one or more pairs of oligonucleotides.

Claims 16 and 17 correspond to claims 5 and 7 as originally filed.

Claims 18 and 19, 21 and 22 refer in a generic form to specific examples 3 and 5. This generalisation is supported by the application documents as a whole and therefore does not add subject-matter.

Claim 20 is based on page 11 disclosing that the size of cells may vary depending on the complexity of the array and that 100 microns is a comfortable upper limit, whereas 10 microns may also be possible to achieve.

2. The insertion of the term "glass" limits the support materials that can be used in the claimed method and therefore results in a restriction of the claimed subject-matter vis-à-vis the subject-matter of the claims as granted.

3. Hence, the claims fulfil the requirements of Articles 123(2) and (3) EPC.

Clarity

4. The appellants argue that the term "... a glass support, to a smooth impermeable surface of which is attached..." not only encompasses the direct, covalent attachment of the oligonucleotides to the glass surface, but also their indirect attachment through a second layer situated on top of the glass surface. Since this layer is only defined by the two relative terms "smooth" and "impermeable", it is impossible to exactly define its nature and consequently the claim lacks clarity.

4.1 In the board's view, the comma after the term "support" in connection with the reference "of which" in the expression after the comma "to a smooth impermeable surface of which" makes it clear that the definition provided by the expression after the comma refers to the surface of the glass support itself.

4.2 Moreover, support for this interpretation comes from the disclosure of the patent as a whole, which pursuant to Article 69 EPC must be taken into account in order to arrive at a technically sensible interpretation of a claim. At no place the patent in suit discloses a second layer and although in all examples the oligonucleotides are immobilized on a microscope slide which is chemically modified with an aliphatic linker, this modification does not create a second surface in addition to the surface of the glass support.

4.3 Thus, the board considers that in view of points 4.1 and 4.2 above, the second interpretation, i.e. the presence of a second, different material layer, is ruled out.

5. The appellants further argue that even if the terms "smooth" and "impermeable" are considered to define the direct surface of the glass support, such a definition of the surface is ambiguous, because "smooth" is a relative term.

5.1 Relative terms constitute a potentially unclear element due to their characteristic to change their meaning according to the context. In the case law such terms were nevertheless considered as clear and their use in a patent therefore allowed, if their meaning was clear in the context of the whole disclosure. This was the case, for example, in the following decisions and for the following terms: T 860/93 (OJ EPO 1995, page 47) - "water-soluble"; T 860/95 of 27 October 1999 - "a long period of time"; T 649/97 of 8 December 2000 - "transparent"; T 1041/98 of 22 October 2001 - "thin plate"; T 193/01 of 4 June 2004 - "thin film composite".

In contrast, in decision T 728/98 of 12 May 2000 the term "substantially pure" was considered unclear per se and in the light of the description.

5.2 The board agrees that "smooth" is a relative term. Thus, the question is whether its meaning is clear in the context of the patent in suit.

As stated above, in all the examples the glass support is represented by a microscope slide whereas the attachment of oligonucleotides to controlled pore glass (part 5.3 of the patent) is marked "for reference purposes". Hence, the frame for the degree of smoothness is delimited: The glass surface is smoother than controlled pore glass and as smooth as glass of ordinary microscope slides or smoother.

Thus, in the board's view, in the context of the disclosure of the patent in suit as a whole, the term "smooth" represents a clear definition of which form the glass support may take and does therefore not render the claim unclear.

6. Hence, the claims fulfil the requirement of Article 84 EPC.

Sufficiency of disclosure

7. According to the case law of the Boards of Appeal the disclosure of a patent must allow the skilled person to perform the invention over the whole range claimed without undue burden (for example decisions T 409/91, OJ EPO 1994, page 653 or T 694/92, OJ EPO 1998, page 97).

8. Doubts as to whether an invention can be carried out in the whole claimed area must according to decision T 19/90 (OJ EPO 1990, page 476) be substantiated by verifiable facts.

9. The method according to claim 1 comprises the covalent attachment of oligonucleotides to different cells on a surface. The cells are arranged in the form of a rectangular matrix, i.e. an array. This arrangement permits any spot to be readily identified by reference to coordinates giving the row and column number of the particular spot. Claim 1 does not contain a restriction as to the number of spots per surface area. Hence, it covers methods for analysing polynucleotides at a high density of nucleic acid spots.

9.1 The appellants argue that the patent in suit does not enable the generation of such arrays because automatic equipment for their preparation was either not available at all at the priority date of the patent in suit or, at least, its preparation amounted to an undue burden. In order to substantiate their argument they rely on a statement in part 5.2 of the patent in suit saying that "automatic equipment for applying the precursors has yet to be developed", on a statement on page 8 of the post-published document OD32 that "at the present time there is a need for methods of nucleic acid sequence analysis which can be automated so that they can be applied on a large scale" and on a statement on page 1008 of post-published documentOD33 that "it is not easy to synthesize large numbers of oligonucleotides even with automated methods".

9.2 The board is not convinced by this line of argumentation. Stating, as in document OD32, that there is a need for a certain method does not necessarily mean that no other methods existed before. Likewise, saying as in document OD33 that something is not easy to carry out does not mean that it is impossible. And indeed, the patent in suit itself demonstrates in Example 5 that a pen plotter can be adapted to deliver nucleotides to a glass surface for synthesis of oligonucleotides and how that can be done: "The pen of the plotter had been replaced by a component, fabricated from Nylon, which had the same shape and dimensions as the pen, but which carried a polytetrafluoroehylene (PTFE) tube, through which the chemicals could be delivered to the surface of the glass slide which lay on the bed of the plotter: A microcomputer was used to control the plotter and the syringe pump which delivered the chemicals. The pen, carrying the delivery tube from the syringe, was moved into position above the slide, the pen was lowered and the pump activated to lay down the coupling solution. Filling the pen successively with G, T and A phosphoramidite solutions an array of twelve spots was laid down in three groups of four, with the different oligonucleotide sequences." Hence, the board concludes that automatic equipment could be made.

10. Furthermore, the appellants, by referring to decisions T 994/95 of 18 February 1999, T 188/97 of 08 February 2001, T 412/93 of 21 November 1994 and T 639/95 of 21 January 1998 argue that, according to the case law of the Boards of Appeal, sufficiency of disclosure has to be denied when the total amount of required experimentation is so high as to amount to an

undue burden. They conclude that, even if the automatic equipment could be made, the amount of time and effort needed for its preparation would be so high as to amount to an undue burden.

10.1 An objection of this kind, to the effect that carrying out an invention involves undue burden, must, like an objection that the invention cannot be carried out or cannot be carried out over the whole claimed scope, be substantiated by verifiable facts (see point 8 above). In contrast to the situation in the decisions of the Boards of Appeal cited in point 10 above, the board does not in the present case have at its disposal any convincing evidence of how high the amount of effort needed would be, and can therefore not come to the conclusion that an undue burden is involved.

10.2 Hence, the available evidence does not allow the board to arrive at the judgement that the manufacture of automatic equipment for preparing large arrays amounts to an undue burden. Consequently, the objection of lack of sufficiency of disclosure based on this argument must fail.

11. In a further line of argumentation the appellants submit that the claimed method as such cannot be carried out as it neglects the teaching in post-published document AD100 which discloses a modern method for preparing a high-density array which differs in some aspects from the one disclosed in the patent in suit. However, this argument must also fail. A piece of evidence describing a specific solution to a posed problem cannot be considered to be a proof that a different solution to the same problem is not operable.

Thus, there is no evidence on file that a skilled person, when trying to prepare a high-density array, would fail as a result of deviating from the teaching in document AD100.

12. According to the relevant case law of the Boards of Appeal the question of the allowable width of a claim in relation to sufficiency depends on the evidence on file in each case. Therefore, in some cases the boards found on the basis of the available evidence that the subject-matter of broad claims was not disclosed by the specification in a manner sufficiently clear and complete for it to be carried by a skilled person, for example T 694/92 (supra), whereas in others they found that it was, for example T 412/93 (supra). This view is also expressed by the board in decision T 636/97 of 26 March 1998.

13. The board finds that the appellants in the present case did not provide evidence showing that the invention claimed is not disclosed in a manner sufficiently clear and complete for it to be carried out by a person skilled in the art. Therefore, the requirements of Article 83 EPC are met.

Novelty

14. The appellants consider document OD2 as novelty-destroying to the subject-matter of claims 1, 14 and 15. It is argued that this document implicitly discloses the use of a glass support, to the smooth impermeable surface of which are covalently bound nucleic acids.

15. The implicit disclosure content of a document is the information which the skilled person derives from it directly and unambiguously on the basis of his/her common general knowledge in combination with the explicit disclosure of that document.

16. Thus, at first, the explicit disclosure content of document OD2 has to be determined. Document OD2 relates to a method of analyzing an analyte in a sample including optional control procedures. The compound to be analysed may be a protein or a nucleic acid. The general description of the support is as follows:

"The solid support itself may assume a variety of configurations such as, eg, a flat rectangular sheet; a round sheet; a rod; a stick; a cylinder; etc. Preferably, the support is a flat, rectangular sheet. For obvious reasons the solid support is preferably water-insoluble and flexible. The solid support may be made of a material selected from the group consisting of: polyvinyl, polystyrene, cellulose, nylon or glass (eg, solid, fiber, etc)" (page 6).

"The solid support is insoluble in the solution being analyzed and is preferably sufficiently flexible to provide for ease of manipulation. The support of the invention may assume a variety of configurations and may be round and flat, card-like, e.g., square or rectangular and flat, tubular, a rod, a stick, cylinder, etc. The support used may obviously be any material capable of maintaining its general configuration during use" (page 21-22).

The specific examples describe attachment of immunoglobulins to a polystyrene support. On page 44 it is mentioned that "nucleic acids and polysaccharides binding to plastic surfaces may be facilitated by a polycationic linking polymer such as polylysine".

Claims characterizing the support read: "17. The system as defined by claim 1 wherein said solid support is insoluble and flexible, and is selected from the group consisting of: a flat, rectangular sheet; a round sheet; and rod; a stick; or a cylinder." and "18. The system as defined by claim 17 wherein said support is made of a material selected from the group consisting of: polyvinyl, polystyrene, cellulose, nylon or glass."

16.1 Hence, the skilled person derives from the whole explicit disclosure that the kind of glass contemplated as support in document OD2 should be water-insoluble and flexible.

17. The disclosure in document OD2 could be considered as novelty-destroying if the skilled person on the basis of his/her common general knowledge about immobilising nucleic acids were aware of any water-insoluble and flexible glass that, at the same time, had a smooth impermeable surface.

17.1 The common general knowledge is represented by basic handbooks and textbooks published before the relevant date of the document of which the implicit disclosure content is to be determined. However, such a type of document is not on file.

17.2 The only document in these proceedings published before the priority date of document OD2 and using a "glass" support is document OD15, published in July 1983. It discloses chemical synthesis of oligonucleotides on glass fibre filters. Even if, for the sake of argument, this document is considered to reflect common general knowledge because it was published long before the priority date of document OD2, the board considers that its contents would not have led the skilled person to imply that the reference in document OD2 to "glass" was to glass with a smooth and impermeable surface. The glass fibre filters disclosed in document OD15 fulfil the characteristics of the support described in document OD2, i.e. they are water-insoluble and flexible. They do not, however, comply with the definition in claim 1 because they are at least not impermeable.

17.3 Hence, the board judges that document OD2 does not implicitly disclose the use of a glass support to a smooth impermeable surface of which are attached oligonucleotides in an assay as defined in claim 1.

18. The reasons for finding the subject-matter of claim 1 novel apply as well to the subject-matter of claims 14 and 15 because the glass support as defined in claim 1 is a feature of these claims, too.

19. The subject-matter of the claims fulfils the requirements of Article 54 EPC.

Inventive step

20. The appellants take the view that a lecture held by Prof. Ekins or either of documents OD1 or OD2 are candidates for the closest prior art.

20.1 Prof. Ekins' lecture refers to a microarray format technique in the immunoassay field. Documents OD1 and OD2 both deal with the detection of nucleic acids.

In accordance with the problem and solution approach, the Boards of Appeal have developed in their case law certain criteria for identifying the closest prior art which provides the best starting point for assessing inventive step. It has been repeatedly pointed out that this should be prior art relating to subject-matter conceived for the same purpose or aiming at the same objective as the claimed invention (cf. Case Law of the Boards of Appeal of the European Patent Office, 4th Edition 2001, chapter I.D.3).

20.2 The present invention relates to the field of nucleic acid analysis. Hence, in the board's judgement either of documents OD1 or OD2 and not the lecture of Prof. Ekins relating to protein analysis are appropriate documents.

20.3 If several documents relate to the same purpose, a secondary consideration for the selection of the closest prior art document is the highest degree of identity of technical features (cf. Case Law of the Boards of Appeal of the European Patent Office, 4th Edition 2001, chapter I.D.3).

20.4 In the assay of document OD2 a binding partner for the molecule to be analysed is affixed onto a surface, and the compound to be analysed, the analyte, binds to its fixed partner. The presence of the bound analyte is determined by a labelled molecule, a probe, which is either the analyte itself (competitive assay) or an antibody to the analyte (sandwich assay) (page 2).

20.5 In the method of document OD1 the nucleic acids to be analysed are labelled and the probes are immobilized on the support. The presence of the analyte is determined by binding of the labelled analyte to the bound probe. The method according to claim 1 relies on the same principle. Hence, document OD1 is the closest prior art document.

21. It was not disputed by the parties that document OD1 discloses all features of claim 1 with the exception of the support material to which the probes are attached. In this respect document OD1 states in general terms on page 9: "Useful solid supports are well known in the art and include those which bind nucleic acids either covalently or non-covalently. Non-covalent supports which are generally understood to involve hydrophobic bonding include naturally occurring and synthetic polymeric materials, such as nitrocellulose, derivatized nylon and fluorinated polyhydrocarbons, in a variety of forms such as filters, beads or solid sheets. Covalent binding supports (in the form of filters, beads or solid sheets, just to mention a few) are also useful and comprise materials having chemically reactive groups or groups, such as dichlorotriazine, diazobenzyloxymethyl, and the like, which can be activated for binding to polynucleotides."

The specific supports materials disclosed in the examples are nylon membranes or nitrocellulose.

22. The next step in assessing inventive step in accordance with the problem-solution-approach is the definition of the technical problem to be solved as the object of the invention in order to generate those effects to be achieved by the claimed subject-matter compared with that of the closest state of the art (cf. Case Law of the Boards of Appeal of the European Patent Office, 4th Edition 2001, chapter I.D.2).

22.1 The problem to be solved is defined differently by the parties. While the appellants consider it to be the provision of an alternative method of analysing a polynucleotide sequence, the respondent argues that it was the provision of an improved method.

22.2 According to the case law of the Boards of Appeal advantageous effects may only be taken into consideration, when formulating the problem underlying an invention, if they are supported by a comparison with the closest prior art (cf. Case Law of the Boards of Appeal of the European Patent Office, 4th Edition 2001, chapter I.D.4.4). Such comparative evidence with regard to document OD1 is lacking here. Hence, the board sees the problem as the provision of an alternative method for parallel analysis of nucleic acid by hybridisation and wherein the nucleic acid is covalently bound to the support.

23. The invention as claimed in claim 1 solves this problem by providing a process as disclosed in document OD1, but using an alternative support material, i.e. nucleic acid is attached to the smooth impermeable surface of a glass support.

Examples 1, 2, 3 and 5 of the patent in suit demonstrate synthesis of oligonucleotides on the surface of a microscope slide resulting in covalent attachment. Specific hybridisation of labelled polynucleotides with the bound oligonucleotides is detected.

Hence, the patent in suit demonstrates that the proposed alternative solution is suitable to solve the problem.

24. The next question to be considered here is whether the skilled person at the priority date of the patent in suit when seeking to solve the problem underlying the invention would have been led in an obvious manner by the closest prior art document OD1 or other documents on file to progress from the closest prior art to the proposed solution.

25. No suggestion is found in the closest prior art document OD1 to seek for an alternative support material, let alone glass with a smooth impermeable surface.

26. The appellants argue that the replacement of the materials disclosed in document OD1 by glass with a smooth impermeable surface is obvious in view of a

combination of document OD1 with either of documents OD13 or OD61.

27. Document OD13 is a patent application relating to the synthesis of oligonucleotides onto a polymeric support. "Glass" as support is mentioned once in the document on page 11: "A wide range of polymer supports can be used as the polymeric support of the present invention. The preferred polymer supports include polystyrenes, crosslinked polystyrenes, cross-linked polyamino acids, polyethyleneglycol, co-polymers of vinyl acetate and N-vinyl pyrrolidone, as well as other polyolefins, polyesters, polyamides, polyacrylates, polymethacrylates, metal oxides, clays, **various glasses** and grafts using combinations of any of these supports."

Document OD61, a scientific publication, discloses covalent attachment of oligonucleotides to long-chain alkylamine and carboxyl controlled pore glass having a pore size of 500 Angström and a diameter of 125µ to 177µ.

The appellants argue that a skilled person would have recognized that the coupling chemistry disclosed in either of documents OD13 or OD61 is suitable for attachment of oligonucleotides to the smooth impermeable surface of a glass support and that he/she would have consequently used it in the method of document OD1.

28. The board assumes for the purpose of the following reasoning that the coupling **method** disclosed in documents OD13 and OD61 might have been suited for

immobilizing nucleic acids on glass with a smooth and impermeable surface. But even if this is so, the question to be answered remains nonetheless whether or not the skilled person would have recognized in view of the prior art that glass with a smooth impermeable surface is suited as a support material in the context of the nucleic acid assay disclosed in document OD1.

29. The following documents illustrate support materials that had actually been used for nucleic acid immobilization up to the priority date of the patent in suit in May 1988.

29.1 As stated above, document OD61, published in July 1987, discloses attachment of oligonucleotides to alkylamine and carboxyl controlled pore glass. Document AD95 discloses in its introduction how controlled pore glasses (CPG) are prepared. The description ends with the sentence: "The borate phase is leached out by acid solutions at high temperatures. The remaining glass contains colloidal silica particles, which are removed by a treatment with NaOH followed by washing with water. The final glass has a porosity between 50% and 75%, and an average pore size between 4.5nm and 400nm. CPG has a surface area between 10 and 350 m^2/g, depending on the pore size." Hence, controlled pore glass is a particulate material with a porous surface and therefore, different from glass having a smooth and impermeable surface.

29.2 The same is true for the glass fibre filters disclosed as support for structured arrangement of nucleic acids in document OD15, published in July 1983 (see point 17.2 above).

29.3 In document OD13, published in March 1984, specific examples of supports for immobilizing nucleic acid are a methacrylate polymer, Amberlite CG50, chloromethylstyrene beads, polyacrylmorpholide resin and a teflonwool/copolymer graft. This selection reflects the statement on page 12 that for the purpose of the method disclosed in document OD13 "polymeric supports with large surface areas consisting of a great number of bonding sites in proportion to weight are most preferred."

29.4 Document OD18, a patent application published in February 1988 and dealing with a nucleic acid hybridisation sandwich assay discloses supports based on beads of amino-, sulfhydryl- or carboxyl-derivatized controlled pore glass, dextran or polystyrene.

29.5 Beads as a support are, inter alia, also contemplated in document OD19, published in 1986: "A variety of solid supports are suggested for the immobilised sample single-stranded polynucleotide including activated glass beads, polyacrylamide, agarose or Sephadex beads and cellulose" (page 2), as well as in documents AD96 and AD97, both published in 1987 and disclosing covalent attachment to submicron latex beads.

29.6 Further support materials referred to are: Nylon membranes (OD1, published in 1987), modified or unmodified cellulose filters (OD20, published in 1983; OD50, published in 1985), or nitrocellulose (OD6, published in 1977).

29.7 Document OD38, entitled "Hybridization properties of immobilized nucleic acids" and published in 1987 summarizes as follows: "The nucleic acid hybridization protocol most familiar to molecular biologists involves the detection by radioactively labelled probes of target nucleic acids which have been immobilised on nitrocellulose or nylon filters. [...] During this same period , Gilham (6,7) described a chemistry of immobilization in which oligonucleotide-length DNA was covalently attached to cellulose supports..." The document itself describes a dextran support.

29.8 Hence, the prior art relating to nucleic acid immobilization on supports does not suggest glass with a smooth impermeable surface as support material.

30. On the other hand, glass plates had for a long time been a common support material in the field of immunology. Document OD45, published in 1984 (and the related US patent OD87, published in 1986) states: "The immunoassay device of this invention comprises a support which has on its surface an array of antibody-coated areas or spots. Preferably, the support is a solid substance having a flat, planar surface such as a glass or plastic coverslip."

31. Furthermore, there is a group of patent documents on file relating to binding assays (AD102 to AD104, published in 1979, 1980, 1981, respectively, AD106, published in 1982, AD107, published in 1984) mentioning DNA and antibodies in one breath when the types of binding partners suited for the respective assays are described. These documents seem to suggest to the skilled person that antibodies and nucleic acids are

alternative types of binding agents that can be interchangeably used in assays for their respective binding partners. The appellants conclude therefrom that these documents suggest that a support material which is good for antibodies is also good for nucleic acids.

32. Finally, document OD41, a US patent application published in 1987, relates to the determination of the site-specific redox state of a liquid system by employing a photoresponsive element. The document belongs to the field of electrochemistry. This is most apparent from the Figures showing electrical circuits, a diagrammatic view of the photoresponsive device and a graph of the observed voltage with varying redox compositions. However, the disclosed device is suggested also for analysing biological materials. In column 8 the following statement is found: "One could analyze for DNA or RNA sequences [...]. For example, one could bind probes to a glass surface, [...] ." The glass surface is later concretized as a "slide".

33. Hence the picture painted by the prior art documents on file may be summarized as follows:

i) In documents relating to the field of immunology glass had been disclosed as a support material since 1984 (OD45).

ii) In documents dealing generally with binding assays and which seem to suggest that antibodies and nucleic acids are alternative types of binding agents that can be treated in the same way, glass is referred to as a

support material as early as in 1979 (OD41, AD101 to 104, AD106, AD107).

iii) During the whole time period covered by the documents cited in points i) and ii) above and up to the priority date of the patent in suit in 1988, there is not a single disclosure or mentioning of a glass support with a smooth impermeable surface in the context of nucleic acid hybridization or synthesis.

34. The board agrees that the claimed subject-matter might look simple, since glass with a smooth impermeable surface is a common material available in every laboratory, for example in the form of microscope slides. However, if, as in the present case, none of the many documents relating specifically to the immobilization of nucleic acid on supports points to its use, although the field of nucleic acid hybridisation and synthesis had been - as can be seen from the many documents on file and the long period spanned by their publication dates - an area of active research for a long time before the priority date of the patent in suit and although glass had been used as a support in the neighbouring field of immunology, then the invention may be simple, but nevertheless it is not obvious.

35. It follows that even if the skilled person had recognized that the coupling method disclosed in documents OD13 or OD61 was in principle suited for immobilizing nucleic acid on glass with a smooth and impermeable surface, he/she would nevertheless not have been prompted to use that material as a support in the context of the nucleic acid assay disclosed in

document OD1. Consequently, the subject-matter of claim 1 is not rendered obvious by a combination of document OD1 with either of documents OD13 or OD61.

36. A glass support with a smooth impermeable surface is a feature of all the independent claims. Therefore, the reasoning above applies to claims 10, 14, 15 and 18, too.

37. The subject-matter of the claims meets the requirements of Article 56 EPC.

Order

For these reasons it is decided that:

1. The decision under appeal is set aside.

2. The case is remitted to the department of first instance with the order to maintain the patent as amended in the following version:

description:
pages 3 to 5, 7, 8, 10 to 12 of the patent specification,

pages 2, 6 and 9 received during oral proceedings on 12 October 2005

claims:

No. 1 to 22 received during oral proceedings on 11 October 2005.

Registrar: The Chairman:

P. Cremona M. Wieser